JRIPE

Journal of Research in **Interprofessional Practice and Education**

Vol. 4.2
November 2014

Measuring Behaviour as Outcome of Interprofessional Interventions: Are You Willing to Collaborate or Are You Collaborating?

Siegrid Deutschlander, PhD & Sara Mallinson, PhD

Abstract

Background: This article discusses a recently developed inventory of questionnaires by a former working group of the Canadian Interprofessional Health Collaborative (CIHC) to promote interprofessional (IP) intervention (education and practice) and program evaluation. The classification of questionnaires into six outcome levels revealed an unusually large number focusing on behavioural outcomes. Behavioural outcomes are key measures for evaluating IP interventions, and we decided to further explore the design of questionnaires in this inventory.

Methods and Findings: The data presented in this article are based on a systematic search and review of questionnaires published in peer-reviewed journals between 2000 and 2010 that evaluate outcomes related to interprofessional education and collaborative practice. The review was used to construct an inventory that placed questionnaires into six outcome levels: attitudes; knowledge, skills, and abilities; collaborative behaviour of providers at the workplace; collaboration as part of organizational practice; patient satisfaction with collaboration, and provider satisfaction with collaboration. We took a closer look at the subgroup of measures on collaborative behaviour of providers (Level 3 Outcome).

We found that the questionnaires included in Level 3 measure a range of competencies. The wording of the subscale items was at times difficult to interpret. While some statements can be clearly attributed to measuring behaviour, others could be seen as measuring some sort of attitudes, beliefs, or knowledge. Since subscale items tended to be a combination of indicators, it was difficult to attribute questionnaires to one particular level.

Conclusions: Designing questionnaires for evaluating outcomes from IP interventions is a challenge, especially when behavioural competencies are of interest. We would welcome more attention paid to the input of potential users in questionnaire construction, who are mostly left with little guidance on how to interpret the questions. Triangulation of methods to supplement the current focus on subjective outcome evaluations from IP interventions would also enhance this research.

Keywords: Competency evaluation; Questionnaires; IP education; Collaboration; IP competencies

Background

This article discusses a recently developed inventory of quantitative instruments evaluating dimensions of interprofessional education and practice interventions. The inventory (hereafter CIHC inventory) was developed by a former working group of the Canadian Interprofessional Health Collaborative (CIHC) to promote interprofessional (IP) intervention and program evaluation (available at www.chd.ubc.ca/files/2013/05/CIHC_tools_report_Aug26-2012.pdf) [1]. The evaluation of outcomes is an important component of any IP initiative, yet finding the right questionnaires can be challenging. Over the years, many voices have called for

Journal of Research in Interprofessional Practice and Education (JRIPE)

Vol. 4.2

Corresponding author: Siegrid Deutschlander. Email: Siegrid.Deutschlander @albertahealthservices.ca

more rigorous evaluation of IP education and collaboration. For example, Carpenter and Dickinson argued that IP education must be "rigorously and appropriately evaluated" to clearly demonstrate the processes that underpin effective IP education and the longer-term outcomes that may result [2, p. 61]. Quantitative and validated before-and-after measures have been seen to provide "hard" evidence essentially quantifying changes in attitudes, knowledge, and skills [2,3]. Over a decade ago, Hyer and colleagues argued that many existing questionnaires were not appropriate for evaluation since few came with supporting evidence on their psychometric properties, most were time-consuming to administer, and in many cases they appeared to assess personality characteristics rather than IP effects [4].

Validated questionnaires are usually preferable to new ones because they may avoid unnecessary costs and time for development and testing while also facilitating cross-site comparison [5,6]. A prospective evaluator will find a diverse array of pretested, standardized questionnaires to choose from. As the numbers continue to proliferate, finding the right questionnaire (i.e., one that is focused enough to gather data answering the evaluation question) can be challenging. The CIHC working group was tasked to compile a more current inventory, modifying the original four-level typology as created by Kirkpatrick and Kirkpatrick in 1967 [7]. The four levels of the original Kirkpatrick typology included learner reaction to the intervention; learning, including knowledge, skills, and attitudes; behaviour changes at the workplace; and results for the organization (e.g., productivity, human relations, profits). Other inventories also draw on the Kirkpatrick typology, such as Carpenter and Dickinson's list of eighteen questionnaires of IP education interventions split into seven levels (learners' reactions, interprofessional attitudes, self-perception, attitudes to teamwork, skills in teamwork, behaviour transfer of learning to practice, and team effectiveness) [2]. In their review, Barr and colleagues changed the original Kirkpatrick four-level typology to six levels, including reaction to program, modification of attitudes/perceptions, acquisition of knowledge/skills, behavioural change, change in organizational practice, and benefits to patients/clients [3].

It is important to note that alternative typologies to the Kirkpatrick framework have also been used. An inventory by Heinemann and Zeiss focuses on measures of team performance [8]. They identified and grouped sixty-six questionnaires that measured four aspects of process behaviours (dependency, conflict, cohesion, interdependence) and four aspects of task behaviours (orientation, organization, open data flow, and problem-solving). Behaviour measures are integral to all of these. The well-known CIHC interactive framework (www.cihc.ca) comprises questionnaires from the 20 Interprofessional Education for Patient-Centred Collaborative Practice (IECPCPC) projects (funded by Health Canada from 2003 to 2007) as well as qualitative evaluation methods (available on http://www.cihc.ca). It arranges the questionnaires in broader categories such as "educational system" or "professional system" and "learners" or "patients," with relevant outcomes for each system.

The six outcome levels developed in the CIHC inventory are as follows: Level 1 – attitudes (about other disciplines or working with other professions); Level 2 – knowledge, skills, and abilities (prior to or acquired from IP education and collaborative

practice interventions); Level 3 – collaborative behaviour of providers at the workplace; Level 4 – collaboration as part of organizational practice; Level 5 – patient satisfaction with collaboration; and Level 6 – provider satisfaction with collaboration. These six levels can be commonly found in evaluations of IP practice interventions and are generally regarded as amenable to change [9]. While a large number of questionnaires focus on attitudes, knowledge, and skills (largely because they are thought to indicate willingness to work with and trust others [10] as well as being open to change through new experiences such as IP interventions [9]), our particular analysis is on the questionnaires at Level 3: behaviour. There are two main reasons for this: First, the transfer of knowledge and skills into actual behaviour at the workplace is an important goal of IP interventions, if not the key goal. IP programs ultimately seek to change provider behaviour toward greater collaboration with a positive impact on patient, system, and provider outcomes. Second, our interest was piqued by the fairly high number of questionnaires measuring collaborative behaviour in the workplace (Level 3) as compared to other inventories, which yielded none or few questionnaires at this outcome level [2,3].

Behaviour evaluations are often more challenging to design than evaluations of other competencies, including attitudes or knowledge/skills, because of difficulties around definition, observation, and reporting bias [11]. For the CIHC inventory, behaviour was defined as "individuals' transfer of interprofessional learning to their practice" [1, p. 5]. This would imply that an observable action, activity, or performance of a task occurs. When evaluating behaviour, Kirkpatrick and Kirkpatrick recommended that important decisions must be made about when, how often, and how to evaluate [11]. In particular, besides asking people about their behaviour changes after training, co-workers' perceptions of behaviour changes at the workplace should also be solicited. This triangulation of methods is a more comprehensive appraisal of behaviour than solely relying on self-assessments, which are notoriously prone to positive skews, as learners may feel compelled to indicate increases in competencies (knowledge, skills, and behaviours) since it is expected of them [2]. Without assuming a purposeful intention to report more positive outcomes, Eva and Regehr argued that self-assessments are inherently flawed, as they are based on an unguided review of practice and experience [12]. Even individuals in the lowest quartile of performance tend to think of themselves as being above average.

As noted earlier, the evaluation of IP education and practice interventions has greatly evolved over the last decade, generating a myriad of questionnaires. Most questionnaire design and appraisal of their relevance was directed at obtaining strong psychometric properties. Numerous tests are conducted to establish construct validity and internal consistency. These are important considerations, but it is also important to make information about questionnaire content and meaning available to potential users. Understanding how behavioural indicators are designed will add transparency and support informed decision making. Developing useful indicators to measure competencies is difficult at the best of times, and competency statements are often "packaged together in a single sentence" [13, p. 254] with the effect of obscuring what is actually being measured.

With a few exceptions, the CIHC Inventory comprises primarily self-assessment questionnaires relying on the understanding of the respondents to make sense of them. Mallinson argued that psychometric testing may "shed little light on the meaning of the questions and response options to respondents and, therefore, the meaning of respondents' answers" [14, p. 11]. The foundation for a meaningful measurement process requires that appropriate and sensitive measures are selected, focusing on relevant concepts and providing data that are reliable and valid [15]. Therefore, how to interpret questions and statements in a questionnaire may become a challenge for potential users. There are usually no opportunities for discussion or clarification when self-assessment questionnaires are being used to collect data. One hopes that the questions have the same meaning for all respondents, and that this shared understanding extends to the research team to interpret their responses. This challenge of ensuring conceptual clarity and shared meaning between questionnaire developers, implementers, and respondents is one that researchers have been reflecting on for decades [14-17], and yet there is still a tendency to overlook the issue of meaning in the testing and validating stages of questionnaires. Despite having the appearance of "hard" data, the quantitative evaluation of subjective opinions is not definite and unchanging across people and places. The interpretability of data depends on good theoretical foundations, explicit definitions, and robust testing of face validity, as well as a good dose of reflexivity on the part of the evaluators as they apply the questionnaire in a new context. Thus, our interest in examining the questionnaires at Level 3 for behaviour statements was guided by considerations about question design, method triangulation, and respondent feedback.

Methods

The methods section comprises two descriptions: the literature review undertaken for the CIHC inventory and the review process for the questionnaires at Level 3 (behaviour) for this article.

Literature Searches

In August 2012, the former CIHC working group completed a quantitative tools inventory measuring outcomes from IP initiatives with 128 unique evaluation questionnaires as available at the UBC College of Health Disciplines website (www.chd.ubc.ca/files/2013/05/CIHC_tools_report_Aug26-2012.pdf). The search strategy was designed with the assistance of a librarian to capture academic articles on quantitative measurement of IP interventions as published between 2000 and 2010. Articles were selected if they showed the development or modification of a questionnaire (with or without validation), or new empirical applications for an IP intervention in different settings. For modified questionnaires, articles with the original questionnaire were included. Abstracts were excluded if they measured general patient or practitioner satisfaction unrelated to collaborative practice, or if the tool was only specific to the evaluation of a program and considered learner reactions. While a search of the grey literature was not conducted due to resource constraints, reports of projects from the *Interprofessional Education for Collaborative Patient-*

Centred Care (IECPCP) initiative were reviewed for relevant questionnaires. The questionnaires from the IECPCP reports were included in the inventory if they provided additional psychometrics or if they were not previously published.

Considering only English-language abstracts, several searches were done in Medline, Embase, CINAHL, Web of Science, ERIC, and Psych Info. The first search for articles published between January 2000 and October 2009 was followed by a second search in May 2010 to retrieve more recent publications and to retrospectively include the terms "validity" and "psychometrics" from January 2000 onward. The search terms related to interprofessional, patient care teams, psychometrics/questionnaires, evaluation, and quantitative analysis. Two hand searches were also conducted, the first of which reviewed the reference sections of retrieved articles about earlier use(s) of a questionnaire or further methodological details for extraction. The second search reviewed relevant journals including the *Journal of Interprofessional Care*, *Journal of Advanced Nursing*, *Gerontology & Geriatrics Education*, and *Medical Education* from 2000 to 2010.

Abstract review and selection

The abstract review undertaken by the CIHC working group (of which the first author was a member) followed a rigorous process. Prior to the systematic abstract review, interrater reliability among the nine reviewers was established by rating 30 abstracts along a three-point rating scale (highly relevant = the questionnaire fits the six-level outcomes typology; possibly relevant = the questionnaire may fit the six-level outcomes typology but requires further information; or not relevant = the questionnaire does not fit any of the six outcome levels). The discussions following this process identified similarities and differences among individuals' ratings, and were instrumental in developing a consistent abstract review process.

For the main article review, the working group reviewers worked in pairs, each pair reviewing approximately 350 abstracts. The members in the pairs reviewed each others' independent ratings to achieve agreement on pulling articles ("highly relevant" or "possibly relevant") for a complete review of full-text articles. From these articles, information on the questionnaires was extracted according to the categories in the inventory table (subscales, sample, setting, other comments). While one pair extracted the data, another pair reviewed the extractions. Extractions were excluded if both pairs agreed that questionnaires did not meet the inclusion criteria. Overall, 416 articles and reports were retrieved for review, of which 136 met the inclusion criteria. The questionnaires were not rated for quality, psychometric rigor, ease of use, or applicability across contexts, as these factors were difficult to ascertain from the articles. For a more detailed description of the methods, please review the final report on the inventory [1]. Since the working group is no longer operational, the inventory will remain as posted on the website without any further updates (www.chd.ubc.ca/files/2013/05/CIHC_tools_ report_Aug26-2012.pdf).

Approach taken for the questionnaire analysis

For the analysis of the questionnaire content in this article, the authors reviewed the

articles for the description of questionnaires. The overall summary (Table 2) is based on the questionnaire characteristics in the inventory, including type of respondent, as well as patients, psychometrics, and settings by outcome levels for all questionnaires. The review of the questionnaires in Level 3 showed that out of 38, 13 studies (34%) included the complete or partial questionnaire and were reviewed for this discussion [18-30]. Articles without questionnaires were not included in this analysis. We were interested in how scale items were worded and rated to measure different types of behaviour or other aspects of collaboration. Given that there is no standardized language to describe the subscales, reviewing the actual subscale items and the rating scales was key to understanding whether the subscales measure perceptions of behaviour or other types of outcomes (e.g., attitudes, knowledge, or skills). The articles were also reviewed for information on triangulation and qualitative questionnaire validation.

Results: The CIHC inventory

Number of questionnaires by outcome level

The final report on the CIHC inventory includes a 46-page table of questionnaires as described in the 136 reviewed articles [1]. In total, 128 unique questionnaires were identified as relevant to interprofessional education or collaborative practice under the six different outcome levels. While there were 128 unique questionnaires, some were classified under more than one level since they measured competencies under different levels amounting to 146 questionnaires (see Table 1). For instance, different subscales in one questionnaire may measure attitudes (Level 1) and skills (Level 2). Fifty-five articles included the complete or partial questionnaires. Some articles were outside of the specified timeframe starting in 2000 but were included because they showed the original questionnaire or contributed in other ways to questionnaire development (e.g., psychometric properties, new setting, or sample). While a few of the tools were used in numerous studies that modified or retested them with different subscales and settings, the inventory was intended as a comprehensive list of tools rather than a comprehensive list of every article using the tools.

Table 1

Number of questionnaires by outcome level

Level	Content	Number of questionnaires
Level 1	Attitudes	64
Level 2	Knowledge, skills, abilities	20
Level 3	Behaviour	34
Level 4	Organizational level	6
Level 5	Patient satisfaction	8
Level 6	Provider satisfaction	14
	Total number of questionnaires	**146**

Outcome Level 1 (attitudes) comprises half of the questionnaires in this inventory, whereas only six questionnaires measure organizational collaboration (Level 4). The questionnaires themselves are not included in the inventory for copyright reasons, as they require permission from some questionnaire developers for their use. Under separate columns, the inventory lists a number of questionnaire characteristics, including subscale descriptions, settings and samples, psychometrics, and other comments (e.g., references to original questionnaires for modified instruments, contact information on authors, inclusion of questionnaires in article).

Characteristics of the questionnaires

This is a brief overview of questionnaire characteristics as they appear in the inventory, which can be consulted by the readers for more detail [1]. Table 2 shows some of the characteristics of the questionnaires, including the type of respondent, as well as patients, psychometrics, and settings by outcome levels for all questionnaires. If a questionnaire was administered to staff and students in separate studies, it was counted for postlicensure and prelicensure learners. There were 93 questionnaires directed at postlicensure learners, mostly staff at practice settings, but also faculty members if the setting was academic. Fifty-eight questionnaires were used with prelicensure learners. These questionnaires were used with students in academic as well as community settings (for example, while doing a practicum in the workplace). The eight questionnaires for patients/families were grouped separately, asking for patient/family satisfaction with collaborative care. A questionnaire was recorded as validated if at least one study validated it. Across the six outcome levels, 103 validated questionnaires were recorded as compared to 43 unvalidated questionnaires. The most frequent psychometrics reported were factor analysis for concept validation and Cronbach's alpha for internal consistency [30]. Fifty-nine questionnaires were used with participants in academic settings, mostly with students but also with faculty. Ninety-eight questionnaires were used in community settings with staff, students, and patients/families. This summary shows that questionnaires are available for all types of learners and settings, hopefully enabling evaluators to find the right one for their purposes.

Table 2

Tool characteristics by outcome levels

Characteristics	Level 1	Level 2	Level 3	Level 4	Level 5	Level 6	Totals
1. Type of learner							
Postlicensure	36	11	28	5	0	13	93
Prelicensure	38	10	6	1	0	3	58
2. Patients	0	0	0	0	8	0	8
3. Pyschometrics							
Various measures	40	12	31	4	6	10	103
Not reported	25	7	3	2	2	4	43
4. Settings							
Academic	37	11	7	2	0	2	59
Community	33	13	27	5	8	12	98

Discussion of subscale questionnaires for level 3: Behaviour

Level 3 was aimed at questionnaires measuring the perception of current or learned behaviour around collaborative practice (also called teamwork by authors) at the workplace. This level comprised 34 questionnaires (with 31 of these validated), which was the second-largest number of questionnaires in this inventory. With the exception of four questionnaires that were used for observations to record the behaviour of others [20,21,29,30], the remaining questionnaires were designed for self-assessment. Way et al. defined collaborative practice as "an inter-professional process of communication and decision making that enables the separate and shared knowledge and skills of healthcare providers to synergistically influence the client/patient care provided" [32, p. 3]. In a similar way, teamwork was seen as self-regulatory and flexible, as requirements demanded potentially measuring roles and responsibilities, interactions and communication patterns, role interdependencies, and information sources [24]. While collaboration involves attitudes, knowledge, and skills, we expected that the questionnaires at this level should "conceptualize teamwork at the level of observable behaviours and processes" [24, p. 277]. "Behavioural markers" for teamwork comprised "non-technical behaviours that contribute to superior or substandard performance within a work environment" [20, p. 581]. They provide "a common language for discussing non-technical skills and can function as frameworks to structure teaching and debriefing," rather than reliance on "gut feeling" [20, p. 581]. Clearly, behaviour measures of teamwork or collaboration have been of interest to evaluators.

Qualitative testing of questionnaires

All reviewed studies conducted factor analysis and internal consistency testing to ascertain construct validity for their questionnaires. While the testing of questionnaires tends to be a quantitative exercise, four studies also involved experts to develop clearly formulated subscales [19,20,24,28]. For example, Farrell and colleagues conducted three rounds of questionnaire testing with professionally diverse respondents for ten item characteristics (e.g., clarity, simplicity, time and effort to complete instrument, scaling system, and redundancy of items in the instrument) [19]. Fletcher and colleagues conducted "cognitive task analysis interviews" [20, p. 581] to isolate non-technical skills considered important for anesthesiologists. Millward and Jeffries involved focus groups with experts in the field for scale construction [24]. Schroder and colleagues invited respondents to provide suggestions on phrasing of items, defining members of teams, and listing items under different headings [28]. These qualitative methods add important value to quantitative measures in construct design.

Triangulation

Two studies used triangulation with self-assessment questionnaires and outsider evaluations to assess team performance [24,26]. Besides self-assessed behaviours by team members, Quoidbach and Hansenne involved supervisors to rate their teams by indicating their level of agreement with 15 items of a job performance question-

naire [26]. They completed the questionnaire with statements: "I often receive positive feedback from patients or visitors about this team" and "the team makes very few efforts to collaborate with other departments" [26, p. 26]. Millward and Jeffries obtained independent measures of team performance from service managers and training personnel with considerable knowledge of the team [24].

Observational questionnaires

Four studies developed questionnaires for observing team behaviours by trained raters. Three of these studies used simulation training to demonstrate learned interprofessional behaviours [20,21,30], while one study videotaped team interactions at team conferences [29]. Two of these were directed at anesthesiologists to function as part of an interprofessional team [20,21], the third one was directed at a medical emergency team, and the fourth one examined professionally diverse teams in rheumatology [29]. The authors used various labels for the observed behaviours, making it difficult to compare across studies. Commonly measured behaviours by the three studies appeared to comprise leadership [20,21,30], communication [20,21,30], understanding of the situation surrounding a patient case [20,21,30], and anticipation or planning [21,30]. Some specific reported behaviours included group climate [21], utilization of resources/information [30], workload distribution [21], vigilance [21], reevaluation [21], recognition of limitations [30], and decision-making [20], even though the last one may have been included by the others under a different subheading (e.g., inquiry/assertion as measured by Gaba) [21]. It was not clear how "group climate" and "workload distribution" were operationalized since the actual statements were not included in the article [21]. Fletcher's study showed a rater comment for "planning & preparing" as "discussed positioning of patient with surgeon, explained all the intricacies of plan to assistant and trainee" [20, p. 584]. While having overall satisfactory interrater reliability, two studies reported that cognitive skills (e.g., situation awareness, knowledge of the environment, anticipation of and planning for potential problems) were more difficult to rate than behaviours and resulted in lower interrater reliability [20,30]. These studies showed that rating cognitive competencies other than expressed behaviour is more difficult to do. Verhoef et al. scored frequency of behaviour as actual time measured in seconds [29].

Self-assessment questionnaires

Besides the observational questionnaires, Level 3 comprised another nine studies with the complete or partial questionnaire included. These questionnaires measured the self-perception of respondents on their own or other team members' types of collaboration. Table 3 shows these questionnaires with examples of statements that can be considered behavioural (see definitions above) or cognitive statements. Six of these questionnaires comprised subscale items that were operationalized as behavioural and cognitive statements [18,22,23,25,27,28]. Two questionnaires comprised cognition statements only [24,26], whereas one questionnaire focused only on behavioural statements [19].

Table 3

Behavioural and cognitive subscale items and examples as measured in self-assessment questionnaires

Behaviour statements		**Cognition statements**	
Relational coordination scale - [22]			
Communication	"How frequently/timely/ accurately do you communicate with these providers about joint replacement patients?"	Shared goals	"How much do these care providers share your goals for the care of joint replacement patients?"
Problem-solving	"Do these care providers work with you to solve the problem?"	Shared knowledge	"How much do these care providers know about your role in caring for joint replacement patients?"
Rating: Frequently/timely/accurately on 5-point Likert scales			
Index of interdisciplinary collaboration - [25]			
Interdependence and flexibility	"I utilize other professionals for their particular expertise." "My colleagues from other professional disciplines and I rarely communicate."	Interdependence and flexibility	"I am not willing to sacrifice a degree of autonomy to support cooperative problem solving."
Reflection on process	"My colleagues from other professions and I do not evaluate our work together." "My non-social work colleagues are as likely as I am to address obstacles to our successful collaboration."	Newly created professional activities	"I am willing to take on tasks outside of my job description when that seems important."
Shared goals	"When colleagues from different disciplines make decisions together they go through a process of examining alternatives."		
Rating: Not provided			
Medical team training questionnaire - [23]			
Communication	"Workload and task distribution are clearly communicated in our work environment."	Teamwork	"Morale on our team is high." "Our team members have mutual respect for each other." "Our team has a shared vision of how to improve."
		Organizational culture	"Our team views problems from a systems perspective rather than as 'someone's fault'." "I am comfortable intervening in a procedure if I have concerns about what is occurring."

Table 3 (continued)

Behaviour statements		**Cognition statements**	
		Human factor awareness	"When I am interrupted, my patient's safety is not affected." "Nurses should not question decisions made by attending physicians.""Fatigue does not affect my ability to perform my work tasks effectively."
Rating: Level of agreement on 5-point Likert scales			
Group decision-making questionnaire - [27]			
Personal task participation	"Did you ask others about their ideas and opinions?""Did you provide information about the situation/opportunities of the patient?"	Emotions	"Did you feel frustrated and tense about other people's behaviour
		Satisfaction	"How satisfied are you with the quality of the treatment plan?"
Negative socio-emotional behaviour	"Did others express a negative opinion about your behaviour?"	Commitment	"To what extent do you feel committed to the treatment plan?"
Rating: Mixed (agreement, satisfaction, semantic differential) on 5-point Likert scales			
Collaborative Practice Assessment Tool (CPAT) - [28]			
Conflict management	"Disagreements among team members are ignored or avoided."	Mission/ meaningful purpose/goals	"Our team's goals are clear, useful and appropriate to my practice." "There is a real desire among team members to work collaboratively."
Patient involvement in their teams	"Team members encourage patients/clients to be active participants in care decisions."	General relationships	"Respect among team members improves with our ability to work together.""Socializing together enhances team work effectiveness."
		Team leadership	"Team leadership supports interprofessional development.""Our team has a process for peer review."
		General role responsibilities/ autonomy	"Each team member shares accountability for team decisions and outcomes.""Team members acknowledge the aspects of care where members of my profession have more skills and expertise."

Table 3 (continued)

Behavour statements	**Cognition statements**	
	Communication and information exchange	e.g., "I trust the accuracy of information reported among team members."
	Community linkages and coordination of care	e.g., "Our team has established partnerships with community organizations to support better patient/client outcomes."
	Decision-making and conflict management	e.g., "Processes are in place to quickly identify and respond to a problem."
	Patient involvement for their teams	e.g., "The patient/client is considered a member of their health care team."
Rating: Level of agreement on 7-point Likert scales		
Team survey - [24]		
	Team identity	"I am proud to belong to this team.""I feel I have strong ties with other members of this team.""When the team succeeds, I feel pleased for all the members."
	Team potency	"This team has little confidence in itself.""This team has the capability to work well together."
	Shared mental models	"The other team members understand my role in the team.""All team members are aware of where to go for information when they need it.""The team takes into consideration the capabilities of its members.""I am well aware of other team member's skills and abilities."
	Meta-cognition	"All team members are clear about the overall goals of the team.""The team is clear about how it contributes to the overall business."
Rating: Not provided		
Team performance questionnaire - [26]		
	Job satisfaction	"I have to the opportunity to participate in training and learning programs in my team.""I feel valorized in my work."
Rating: 17 items with 5-point Likert scales		

Behaviour statements		**Cognition statements**	
Team climate inventory - [18]			
Team member interactions in task orientation	"Do you and your colleagues monitor each other so as to maintain a high standard of work?""We keep in touch with each other as a team."	Shared vision	"How worthwhile do you think these objectives are to your team?""To what extent do you think these objectives are realistic and can be attained?"
Support for innovation	"People in this team are always searching for fresh, new ways of looking at problems.""This team is always moving toward the development of new answers."	Participa nt safety	"We all influence each other.""People feel understood and accepted by each other."
Rating: Mixed (agreement, extent of interactions, truth value of statements, extent of perceived safety)			

Table 3 (continued)

Behaviour statements		**Cognition statements**
Medication Use Process Matrix (MUPM) - [19]		
Diagnosing and prescribing	Make the diagnosis, select best drug for patient.	
Monitoring	Monitor effectiveness and safety, monitor compliance, receive and organize requests for prescription renewals.	
Administrative/ documentation	Complete forms as required, manage drug samples.	
Education	Educate patients about the medication; provide group patient education regarding medications.	
Medication review	Identify prescribing errors, screen patients' medication lists, provide complete medication review.	
Rating: Level of contribution on 5-point Likert scales		

Some authors indicated in their description of the study and questionnaire development what type of measures they focused on to assess collaboration in the workplace. For example, the validated Relational Coordination Scale developed by Hoffer Gittell [21] measures various coordinating tasks among team members for joint-replacement patients. In a similar way, Parker-Oliver and colleagues reported on the perception of team members on a series of activities by type of team [25]. Mills and colleagues modified a previously developed team training questionnaire "to elicit more specific information related to communication and teamwork between clinicians" in surgical teams [23, p. 107]. Roelofsen et al. adapted a questionnaire on "group decision-making processes" at team conferences for rehabilitation teams [27, p. 149]. Schroder et al. used the Collaborative Practice Assessment Tool (CPAT) to measure the perceptions of team members on "working collaboratively" [28, p. 189]. The authors of one study explicitly stated their interest in measuring the cognitive dimensions influencing team processes [24]. Subscales and subscale items varied across questionnaires since the authors based their questionnaires on different theories for and definitions of collaboration. However, given the focus on collaborative activities and teamwork, we expected most of the questionnaires at this level to include some behavioural statements in their subscale measures.

Behaviour measures

The seven questionnaires with behaviour statements or questions measured common aspects of collaboration under different subheadings. For instance, communication-related items (also included under personal task participation and interdependence and flexibility) include "how frequently/timely/accurately do you communicate with these providers about joint replacement patients?" [22, p. 3], "did you provide information about the situation/opportunities of the patient?" and "did you ask others about their ideas and opinions?" [27, p. 154], "my colleagues from

other professional disciplines and I rarely communicate" [25, p. 283], and "workload and task distribution are clearly communicated in our work environment" [23, p. 109]. Anderson and West show communication as "we keep in touch with each other as a team" [18, p. 24]. While these are action statements, another item for communication was included under cognitive measures since it expressed some form of trust in current behaviour ("I trust the accuracy of information reported among team members") [28, p. 4].

A few statements related to problem-solving or conflict management, albeit under different headings (e.g., support for innovation). The following statements could be included here: "did these care providers work with you to solve the problem?" [22, p. 26], "my non-social work colleagues are as likely as I am to address obstacles to our successful collaboration" [25, p. 283], "disagreements among team members are ignored or avoided" [28, p. 5], and "people in this team are always searching for fresh, new ways of looking at problems" [18, p. 247]. The Medication Use Process Matrix also included medication review for identifying errors (e.g., identify prescribing errors, screen patients' medication lists, provide complete medication review) [19].

Other behavioural statements are reflections on process or monitoring: "my colleagues from other professions and I do not evaluate our work together" [25, p. 284], "do you and your colleagues monitor each other so as to maintain a high standard of work?" [18, p. 247]. These statements are examples of monitoring various aspects of the medication review in the Medication Use Process Matrix (e.g., monitor effectiveness and safety, monitor compliance, receive and organize requests for prescription renewals) [19].

Miscellaneous other types of behaviours were captured in these questionnaires. Schroder et al. described patient involvement in teams as "team members encourage patients/clients to be active participants in care decisions" [28, p. 5]. Besides this behavioural statement, the authors also expressed patient involvement in teams as an attitude that can lead to active participation of patients in their care, as in "the patient/client is considered a member of their health care team" [28, p. 5]. Anderson and West wrote "this team is always moving toward the development of new answers" [18, p. 247] as an indicator for innovation. An item that could have been easily worded as an emotional statement was this one provided by Roelofsen et al.: "Did others express a negative opinion about your behaviour?" [27, p. 154]. Given that the statement specifically asked for expressing opinions, it referred to action rather than emotion/attitude.

Cognitive, attitudinal, and emotional measures

Eight self-assessment questionnaires measure dimensions of collaboration that differ from behaviour [16,18,19,22,23,25,27,28]. These include attitudes, knowledge, skills, and/or intentions, depending on the wording of the statements.

Some questionnaires include aspects of collaboration related to shared goals, a common vision, or meaningful purpose. Hoffer Gittell developed the statement "how much do these care providers share your goals for the care of joint replacement patients?" [22, p. 19]. Schroder and colleagues included the items "our team's goals are

clear, useful and appropriate to my practice" and "there is a real desire among team members to work collaboratively" [28, p. 2]. Mills et al. included a similar statement ("our team has a shared vision of how to improve" [23, p. 109]) placed under teamwork . These statements are phrased as positive value statements that may enhance the likelihood for collaboration.

Various other statements cluster around teamwork climate and interpersonal relationships, expressing attitudes toward team cohesion. Mills and colleagues included "morale on our team is high" and "our team members have mutual respect for each other" [23, p. 109] in the subscale teamwork. Parker-Oliver referred to "I am not willing to sacrifice a degree of autonomy to support cooperative problem solving" [25, p. 283] under independence and flexibility. Schroder et al. offered statements under general relationships, including "respect among team members improves with our ability to work together," "socializing together enhances team work effectiveness," and under team leadership, "our team has a process for peer review" [28, p. 213].

Two questionnaires included provider satisfaction with various team aspects or workplace features [26,27]. Quoidbach and Hansenne referred to aspects of job satisfaction as "I have the opportunity to participate in training and learning programs in my team" and "I feel valorized in my work" [26, p. 25]. Roelofsen and colleagues asked respondents to express their satisfaction with the quality of the treatment plan [27]. Given that satisfaction questionnaires are classified at a different level of the CIHC inventory (Level 6), these questionnaires should also be cross-referenced with that level.

In some cases, subheadings surprised us by providing attitudinal or emotional instead of behavioural statements. For example, Parker-Oliver et al. included a statement expressing a level of "willingness" for newly created professional activities ("I am willing to take on tasks outside of my job description when that seems important"), which could have been equally worded as a clear behavioural statement (I take on tasks outside of my job description) [25, p. 283]. Merely indicating "willingness" for taking action indicates intention or an attitude but cannot be seen as actual behaviour. For Mills et al., "teamwork" statements including "morale on our team is high," "our team members have mutual respect for each other," and "our team has a shared vision of how to improve" [23, p. 109] are more oriented toward common views and attitudes. Schroder et al. included under decision-making and conflict management "processes are in place to quickly identify and respond to a problem" and under patient involvement for their teams "the patient/client is considered a member of their health care team" [28, pp. 2, 3]. Both statements hint at the potential behaviour of team members to follow their procedures or actually involve patients but do not actually measure it.

Some questionnaires had similar subheadings but mixed behavioural and cognitive statements. For example, shared goals were measured by two studies [22,25]. Parker-Oliver and colleagues operationalized shared goals as a behavioural statement ("when colleagues from different disciplines make decisions together they go through a process of examining alternatives" [25, p. 285]), whereas Hoffer Gittell phrased it more as a cognitive statement ("how much do these care providers share your goals for the care of joint replacement patients?" [22, p. 19]). Parker-Oliver and

colleagues also operationalized items under the same subheading, "interdependence and flexibility," in two different ways [25]. Their behavioural statements for this dimension read as "I utilize other professionals for their particular expertise" and "my colleagues from other professional disciplines and I rarely communicate" [25, p. 283]. The other statement for this dimension expresses a level of "willingness," as in "I am not willing to sacrifice a degree of autonomy to support cooperative problem solving" [25, p. 283].

Miscellaneous other statements were included in the questionnaires. Another emotional statement asks "did you feel frustrated and tense about other people's behaviour?" [27, p. 154]. For Hoffer Gittell, shared knowledge is also part of the Relational Coordination Scale, as in "how much do these care providers know about your role in caring for joint replacement patients?" [22, p. 18]. While emotions could be subsumed under attitudes (Level 1), knowledge questions fit Level 2.

Conclusions

Our analysis of the self-assessment questionnaires with complete or partial tools indicates that most of the tools grouped under Level 3 did not exclusively measure behavioural competencies. The Level 3 questionnaires showed that most studies aimed at assessing the perception of behaviour as well as cognitive dimensions, including attitudes, awareness, and satisfaction. As used by most researchers, the definition of collaboration is broad, comprising behavioural as well as cognitive dimensions of collaboration related to attitudes, beliefs, values, satisfaction, and knowledge. Beside the four observation studies, out of nine self-assessment questionnaires, only the Medication Use Process Matrix measured exclusively behavioural competencies relating to medication management and dispensing [19]. Given that the typology lists attitudes, knowledge, skills, and behaviour at different levels, eight of these questionnaires should be cross-referenced with the other levels too. Without having done a detailed examination, we would expect similar results emerging from the questionnaires at the other levels.

Our subscale analysis indicates that questionnaire assignment to the different outcome levels was less clear than expected. While authors may use different theories and therefore include different dimensions of collaboration, it becomes a methodological issue to translate these dimensions into clearly operationalized indicators. In some cases, it was not obvious whether obtaining behaviour measures was the intent of the questionnaire designers. Clearly, willingness to take certain actions is not the same as actually taking these actions. Was "willingness" for newly created professional activities intentionally chosen over an action statement that could indicate the engagement in professional activities? In other cases, it was not clear whether a statement indicated an attitude or some other emotional dimension. Upon closer examination, the competency statements were truly "packaged together," in some cases obscuring what is supposed to be measured.

The questionnaires also raised the question of whether a modified Kirkpatrick typology was useful for classifying IP questionnaires at all. Kirkpatrick's original typology comprised only four levels with learner reaction, learning, behaviour, and

results [7]. Alternative typologies have been developed comprising outcomes other than those in the Kirkpatrick typology. The subsequent differentiation of "learning" as dimensions of attitudes, knowledge, skills, and satisfaction does not reflect the multi-dimensional nature of questionnaires that combine various dimensions often into one subscale. It also shows how difficult it can get to clearly distinguish one type of measure from another. Kirkpatrick also recommended evaluating behaviour as assessed by others through observations [11]. This would make behaviour evaluations a distinct objective to be assessed through mainly observational methods as compared to self-assessments. The development of self-assessment tools for competency acquisition has therefore been an interesting development. It has occurred contrary to competency evaluation elsewhere in healthcare where the focus was on using objective structured clinical skills examinations (OSCEs) [33]. With few exceptions, the questionnaires focus on self-reported competencies of respondents' perceptions on their own attitudes, knowledge, skills, and behaviours. The inventory indicates that by designing these types of questionnaires, we are likely to get biased perceptions of learners on their own competencies.

Another issue emerging from this inventory relates to the lack of standardized wording to describe the measures. Without the questionnaire included in the article, trying to place questionnaires at the various levels based on subscale labels is too ambiguous. Subscale labels are often obscure, and without any example statements they are almost impossible to assign to any dimension of collaboration. The classification of some of the tools to the behaviour level may therefore have been in error since extensive discussions on assigning the questionnaires to the respective levels did not occur among the CIHC working group members. While semantics is an issue for researchers, it surely would be a challenge for potential respondents to interpret various competency statements in these self-assessment tools. Even if the components considered important for collaboration differed among questionnaires, the language used in statements should be clear and distinct. While great care was taken with concept validation, assessing the psychometric properties of the tools and establishing interrater reliabilities, surprisingly little attention was paid to clarifying the meaning of the subscale items. Only four studies applied qualitative testing to clarify subscale wording, which was not described in any great length. As compared to studies using self-assessment tools, evaluations from observations tend to rely on interrater reliability testing, which involves extensive discussions about the meaning of observations. Thus, meaning construction of questionnaire items evaluating IP education and practice interventions is an area where we would like to see further discussion and clarification.

References

1. CIHC Working Group. (2012). *An inventory of quantitative tools measuring interprofessional education and collaborative practice outcomes. Final report*. URL: (http://www.chd.ubc.ca/files/2013/05/CIHC_tools_report_Aug26-2012.pdf) [September 16,2014].
2. Carpenter, J., & Dickinson, H. (2008). *Interprofessional education and training*. Bristol: The Policy Press.
3. Barr, H., Koppel, I., Reeves, S., Hammick, M., & Freeth, D. (2005). *Effective interprofessional education: Argument, assumption and evidence*. Oxford: Blackwell Publishing.

4. Hyer, K., Fairchild, S., Abraham, I., Mezey, M., & Fulmer, T. (2000). Measuring attitudes related to interdisciplinary training: Revisiting the Heinemann, Schmitt and Farrell attitudes toward health care teams' scale. *Journal of Interprofessional Care, 14*(3), 249–258.
5. Forman, D., & Nyatanga, L. (2001). The process of developing a research questionnaire to measure attitudes to shared learning. *Medical Teacher, 23*(6), 595–598.
6. Freeth, D., Hammick, M., Reeves, S., Koppel, I., & Barr, H. (2005). *Effective interprofessional education: Development, delivery and evaluation.* Oxford: Blackwell Publishing.
7. Kirkpatrick, D.L. (1967). Evaluation of training. In Craig, R., & Bittel, L. (Eds.), *Training and Development Handbook* (pp. 87–112). New York: McGraw-Hill.
8. Heinemann, G.D., & Zeiss, A.M. (Eds.). (2002). *Team performance in health care: Assessment and development.* New York: Kluwer Academic/Plenum Publishers.
9. Lindqvist, S., Duncan, A., Shepstone, L., Watts, F., & Pearce, S. (2005). Development of the Attitudes to Health Professionals Questionnaire (AHPQ): A measure to assess interprofessional attitudes. *Journal of Interprofessional Care, 19*(3), 269–279.
10. Anderson, E., Manek, N., & Davidson, A. (2006). Evaluation of a model for maximizing interprofessional education in an acute hospital. *Journal of Interprofessional Care*, *20*(2), 182–194.
11. Kirkpatrick, D.L., & Kirkpatrick, J.D. (2006). *Evaluating training programs: The four levels*, 3rd edition. San Francisco: BK Publishers.
12. Eva, K., & Regehr, G. (2007). Knowing when to look it up: A new conception of self-assessment ability. *Academic Medicine, 82*(S10), S81–S84.
13. Reeves, S. (2012). The rise and rise of interprofessional competency. *Journal of Interprofessional Care, 26*(4), 253–255.
14. Mallinson, S. (2001). Listening to respondents: A qualitative assessment of the short-form 36 Health Status Questionnaire. *Social Science & Medicine, 54*, 11–21.
15. Cicourel, A.V. (1964). *Method and measurement in sociology*. New York: The Free Press of Glencoe.
16. Pawson, R. (1989). *A measure for measures: A manifesto for empirical sociology*. London: Routledge.
17. Tanur, J.M. (Ed.). (1992). *Questions about questions: Inquiries into the cognitive bases of surveys.* New York: Russell Sage Foundation.
18. Anderson, N.R., & West, M.A. (1998). Measuring climate for work group innovation: Development and validation of the Team Climate Inventory. *Journal of Organizational Behaviour*, *19*(3), 235–258.
19. Farrell, B., Pottie, K., Woodend, K., Yao, V. H., Kennie, N., et al. (2008). Developing a tool to measure contributions to medication-related processes in family practice. *Journal of Interprofessional Care, 22*(1), 17–29.
20. Flechter, G., Flin, R., McGeorge, P., Glavin, R., Maran, N., & Patey, R. (2003). Anaesthetist's non-technical skills (ANTS): Evaluation of a behavioural marker system. *British Journal of Anaesthesiology,* 90, 580–588.
21. Gaba, D.M., Howard, S.K., Flanagan, B., Smith, B.E., Fish, K.J., & Botney, R. (1998). Assessment of clinical performance during simulated crises using both technical and behavioral ratings. *Anesthesiology, 89(1),* 8–18.
22. Hoffer Gittell, J. (2004). *Relational coordination: Recommendations for measurement and analysis*. (working paper), available from: jodyhoffergittell@brandeis.edu.
23. Mills, P., Neily, J., & Dunn, E. (2008). Teamwork and communication in surgical teams: Implications for patient safety. *Journal of the American College of Surgeons, 206*(1), 107–112.
24. Millward, L.J., & Jeffries, N. (2001). The team survey: A tool for health care team development. *Journal of Advanced Nursing, 35*(2), 276–287.
25. Parker-Oliver, D., Bronstein, L.R., & Kurzejski, L. (2005). Examining variables related to successful collaboration on the hospice team. *Health & Social Work, 30*(4), 279–286.
26. Quoidbach, J., & Hansenne, M. (2009). The impact of trait emotional intelligence on nursing team performance and cohesiveness. *Journal of Professional Nursing, 25(1),* 23–29.
27. Roelofsen, E., Lankhorst, G., & Bouter, L. (2001). Translation and adaptation of a questionnaire to assess the group processes of rehabilitation team conferences. *Clinical Rehabilitation, 15*(2), 148–155.
28. Schroder, C., Medves, J., Paterson, M., Byrnes, V., Chapman, C., O'Riordan, A., Pichora, D., & Kelly, C. (2011). Development and pilot testing of the collaborative practice assessment tool. *Journal of Interprofessional Care, 25*(3), 189–195.
29. Verhoef, J., Toussaint, P.J., Putter, H., Zwetsloot-Schonk, J.H.M., & Vlieland, T.P.M.W. (2005). Pilot study of the development of a theory-based instrument to evaluate the communication process during multidisciplinary team conferences in rheumatology. *International Journal of Medical Informatics, 74*(10), 783–790.

30. Wallin, C.J., Meurling, L., Hedman, L., Hedegard, J., & Felländer-Tsai, L. (2007). Target-focused medical emergency team training using a human patient simulator: Effects on behaviour and attitude. *Medical Education, 41*(2), 173–180.
31. Kenaszchuk, C. (2013). An inventory of quantitative tools measuring interprofessional education and collaborative practice outcomes. *Journal of Interprofessional Care, 27*, 101–103.
32. Way, D., Jones, L., & Busing, N. (2000). *Implementation strategies: Collaboration in primary care – family doctors & nurse practitioners delivering shared care.* Toronto, Ontario: College of Family Physicians.
33. Morison, S.L., & Stewart, M.C. (2005). Developing interprofessional assessment. *Learning in Health and Social Care, 4*(4), 192–202.